

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 16:09:13 ; Search time 299.73 Seconds  
(without alignments)  
16.815 Million cell updates/sec

Title: US-09-331-631a-3\_COPY\_74\_116  
Perfect score: 250  
Sequence: 1 NODDPOTDCQCCQRCRCROE.....RQOQYCORCKEICEEEY 43

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues  
Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SP:REMBL\_15:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_mhc:\*  
9: sp\_mhc:\*  
10: sp\_mhc:\*  
11: sp\_mhc:\*  
12: sp\_mhc:\*  
13: sp\_mhc:\*  
14: sp\_mhc:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	250	100.0	666	10	Q9SP14 macadamia i
2	242	96.8	666	10	Q9SP15 macadamia i
3	235	94.0	625	10	Q9SP13 macadamia i
4	115	46.0	593	10	Q9SEW4 juglans reg
5	112	44.8	525	10	Q43358 theobroma c
6	92	36.8	411	5	P91419
7	78	31.2	572	5	Q19594 caenorhabd
8	77.5	31.0	425	5	Q17400 caenorhabd
9	77.5	31.0	600	5	Q17401 caenorhabd
10	77	30.8	810	10	Q9ZWI3 cucurbita m
11	73.5	29.4	242	5	Q19519 caenorhabd
12	73	29.2	111	5	Q19054 caenorhabd
13	72	28.8	330	5	Q18118 caenorhabd
14	72	28.8	339	5	Q9N4R1 caenorhabd
15	71	28.4	388	5	Q16500 caenorhabd
16	71	28.4	388	5	Q16501 caenorhabd
17	71	28.4	438	5	Q16502 caenorhabd
18	71	28.4	445	5	Q16511 caenorhabd
19	70.5	28.2	419	5	Q9NA38 caenorhabd

20	70.5	28.2	420	5	Q9NA61	Q9NA61 caenorhabd
21	69.5	27.8	335	5	Q9NF70	Q9NF70 caenorhabd
22	69.5	27.8	335	5	Q9NF69	Q9NF69 caenorhabd
23	69.5	27.8	709	5	Q9N4R2	Q9N4R2 caenorhabd
24	69	27.6	388	5	Q44606	Q44606 caenorhabd
25	68.5	27.4	378	5	Q27383	Q27383 caenorhabd
26	67.5	27.0	314	5	Q23390	Q23390 caenorhabd
27	67.5	27.0	335	5	Q9N6L3	Q9N6L3 caenorhabd
28	67	26.8	393	10	Q9ZTP0	Q9ZTP0 oryza sativ
29	67	26.8	1513	5	Q17970	Q17970 caenorhabd
30	66.5	26.6	342	4	Q9UK28	Q9UK28 homo sapien
31	66.5	26.6	637	10	Q03678	Q03678 hordeum vul
32	66	26.4	165	5	Q9XV49	Q9XV49 caenorhabd
33	64.5	25.8	654	5	Q17982	Q17982 caenorhabd
34	63.5	25.4	273	5	Q45362	Q45362 caenorhabd
35	63.5	25.4	1306	5	Q77273	Q77273 drosophila
36	63	25.2	191	6	Q9N0L8	Q9N0L8 mactopus eu
37	63	25.2	1129	12	Q9GR71	Q9GR71 kaposi's sa
38	62.5	25.0	356	5	Q17316	Q17316 ceratitis c
39	62	24.8	594	5	Q77337	Q77337 plasmodium
40	61	24.4	539	4	Q9NDA2	Q9NDA2 homo sapien
41	61	24.4	646	5	Q9M4I9	Q9M4I9 drosophila
42	60.5	24.2	269	5	Q9B9J2	Q9B9J2 toxocara ca
43	60	24.0	243	5	Q9N0L9	Q9N0L9 entamoeba d
44	60	24.0	425	5	Q15755	Q15755 dictyostell
45	59.5	23.8	294	5	Q9XUS0	Q9XUS0 caenorhabd

## ALIGNMENTS

```

RESULT 1
Q9SP14 ID Q9SP14 PRELIMINARY: PRT: 666 AA.
AC Q9SP14;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE VICILIN PRECURSOR.
GN AMP2.
OS Macadamia integrifolia (Macadamia nut).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.
OX NCBI_TaxID=60698;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=NUT KERNEL.
RA Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;
RT "A family of antimicrobial peptides is produced by processing of a 7S
RT globulin protein in Macadamia integrifolia.";
DR EMBL: AF161884; AMD54245.1; -
DR HSP: P02853; 2PHL.
DR INTERPRO: IPR001113; -
DR PFAM: PF00546; Seedstore_7s; 1.
SQ SEQUENCE 666 AA: 78243 MW: 0ECA22F8710F8A7B CRC64:

Query Match 100.0%; Score 250; DB 10; Length 666;
Best Local Similarity 100.0%; Pred. No. 3.9e-23;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NODDPOTDCQCCQRCROESGPRQOQYCORCKEICEEEY 43
Db 74 NODDPOTDCQCCQRCROESGPRQOQYCORCKEICEEEY 116

RESULT 2
Q9SP15 ID Q9SP15 PRELIMINARY: PRT: 666 AA.
AC Q9SP15;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

```

DT 01-OCT-2000 (TREMBL:rel. 15, Last annotation update)  
DE VICILIN PRECURSOR.  
CN AMP2.  
OS Macadamia integrifolia (Macadamia nut).  
OC Euxaryota: Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; Eudicotyledons; Proteaceae; Macadamia.  
OX NCBI\_Taxid=60698;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=NUIT KERNEL;  
RA Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;  
RT "A family of antimicrobial peptides is produced by processing of a 7S  
RL globulin protein in Macadamia integrifolia kernels.";  
RL Plant J. 0:0-0(1999).  
DR EMBL; AF161883; AAD54244.1; -.  
DR HSSP; P02853; 2PHL.  
DR INTERPRO: IPR001113; -.  
DR Pfam: PF00546; Seedstore\_7s.1.  
SQ SEQUENCE 666 AA; 78217 MW; C752B884B2DF0224 CRC64;

Query Match	96.8%;	Score 242;	DB 10;	length 666;
Best Local Similarity	95.3%;	Pred. NO. 3.8e-22;		
Matches 41; Conservative	2;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	NQDDPQTDCCQCCRRCRQOESGPRQOQYCQRCKEICEEEEEEY	43
Db	74	NQEDPQTECQCCQRCRCRQOESGPRQOQYCQRCKEICEEEEEEY	116

RESULT	3
Q9SPL3	
ID	Q9SPL3
	PRELIMINARY;
	PRT; 625 AA

DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
DE VICILIN PRECURSOR (FRAGMENT).  
GN AMP2.

OS *Macadamia integrifolia* (Macadamia nut).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta  
 OC Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.  
 OX NCBI\_TaxID=60698;

RN  
 RP  
 RC  
 RA  
 RT  
 RL  
 [1]  
 SEQUENCE FROM N.A.  
 TISSUE=NUJ KERNEL;  
 Marcus J P., Coulter K C., Green J L., Manners J M.,  
 "A family of antimicrobial peptides is produced by processing of a 7S  
 globulin protein in *Macadamia integrifolia*.",  
 Plant J. 0:0-0(1999).

DR EMBL; AF161885; AAD54246.1; -.  
DR HSSP; P02853; 2PHL.  
DR INTERPRO; IPR001113; -.  
DR PFAM; PF00546; Seedstore\_7s; 1.  
FT NON\_TER 1 1  
SQ SEQUENCE 625 AA; 7336 MM; 415802A89D370296 CRC64

	Query Match	94.0%	Score 235;	DB 10;	length 625;
	Best Local Similarity	93.0%;	Pred. No. 2,6e-21;		
	Matches	40;	Conservative	2; Mismatches	1; Indels
Oy	1 NODPQTCQCQCRRCRODESGPRDOYTCORRCKETCEEEEBY	43			
	:				
b	33 NOEDPTECQCQCRRCROESDPDPODYTCORRCKETCEEEEBY	75			
	:				

RESULT	4	
Q9SEW4		
Q9SEW4	PRELIMINARY;	PRT; 593 AA.
AC Q9SEW4;		
DT 01-MAY-2000	(TREMBLrel. 13, Created)	
DT 01-MAY-2000	(TREMBLrel. 13, Last sequence update)	

```

DT      01-OCT2000 (TEMBLrel. 15, last annotation update)
DE      VICILIN-LIKE PROTEIN PRECURSOR (FRAGMENT).
OS      Juglans regia (English walnut).
OC      Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC      Magnoliophyta; Eudicotyledons; core eudicots; Rosidae; eurosids I;
OC      Fagales; Juglandaceae; Juglans.
OX      NCBI_TaxID=51240;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=CV. SUNLAND; TISSUE=SOMATIC EMBRYO LINE;
RA      Theuber S.S., Jarvis K.C., Peterson W.R., Dandekar A.M., Ansari A.A.
RT      "Identification and cloning of a cDNA encoding a vicilin-like protein
RT      Jug r 2, from English walnut kernel (Juglans regia): a major food
RT      allergen."
RL      Submitted (MAY-1998) to the EMBL/Genbank/DBJ databases.
DR      EMBL; AF066055; AAF18269.1; -.
DR      HSSP; P02853; 2PHT.
DR      INTERPRO: IPR001113; -.
DR      Pfam: PF00546; Seedstore_7s; 1.
FT      NON_TER      1      1
CO      SEQUENCE      593 AA; 63990 MW; 9BA127E19B18C0D8 CRC64;

```

Query Match	46.08;	Score 115;	DB 10;	Length 593;
Best Local Similarity	51.28;	Pred. NO. 1.3e-06;		
Matches 21; Conservative	6;	Mismatches 14;	Indels 0;	Gaps 0;

```

Oy      1 NODDPQDTCCOCCORCRCKOESGPRDOOYCORCKEICEEE 411
         |  ||  :|||  ||:|  | |||  || ||:|  ||:|
Db      10 NRPDPREYRQCOEYCRRGCGGCRQOQOCCIRKEERLEDD 500

```

RESULT	5
Q43358	
ID	Q43358
PRELIMINARY;	
PRT;	525 AA

DT	01-NOV-1996	(TREMBlrel. 01, Created)
DT	01-NOV-1996	(TREMBlrel. 01, Last sequence update)
DT	01-OCT-2000	(TREMBlrel. 15, Last annotation update)

DE VICILIN PRECURSOR.  
GN CSV.  
OS *Thebroma cacao* (Cacao).  
OC Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta.  
OC Magnoliophyta: eudicotyledons, core eudicots, Rosidae, eurosids II.  
OC Malvales: Malvaceae: Thebroma.  
OX NCBI\_taxid=3641.

RN [1]  
RP SEQUENCE FROM N.A.  
RQ  
RC TISSUE=LEAVES;  
RX MEDLINE=92288309; PubMed=1600151;  
RA McHenry L., Filtz P.J.;  
RT "Comparison of the structure and nucleotide sequences of vicillin genes  
of cocoa and cotton raise questions about vicillin evolution.";  
RL Plant Mol. Biol. 18:1173-1176(1992).

DR	EMBL:	X62625;	CAA44493.1;	-
DR	EMBL:	X62626;	CAA44494.1;	-
DR	HSSP:	P02853;	2PHL.	
DR	MENDEL:	30919;	Phccg1188;	30919.
DR	INTERPRO:	IPRO0113;	-	
DR	PFAM:	PF00546;	Seedstore_7s;	1.
DR	PRODOM:	PD081059;	-;	1.
KW	Signal.			
FT	Signal.	1	24	POTENTIAL.
FT	CHAIN	25	525	VICILIN.
SO	SEQUENCE	525 AA:	60798 MW;	19114CD5C248950D CRC64

	Query Match	44.8%	Score 112;	DB 10;	Length 525;
	Best Local Similarity	47.5%;	Pred. No. 2.7e-06;		
	Matches 19; Conservative	10;	Mismatches 0;	Gaps 0;	
OY	2 ODDPOTDCOCRCRQDESGPRPOOYCCKRKKEICEENE	41			
	:::         :::::	:        :	:	:  :	

```

Db      78  EELIQR0YQCGGRCOEQ0GGRDEQ0GCRKMEGYKKEE 117

RESULT      6
P91419
ID      P91419      PRELIMINARY:      PRT:      411 AA.
AC      P91419:
DT      01-MAY-1997 (TREMBLrel. 03, Created)
DT      01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT      01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE      CODED FOR BY C. ELEFANS CDNA YK115A6.5.
GN      T01D1.6.
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC      Rhabditidae; Peloiderinae; Caenorhabditis.
OX      NCBI_TaxID=6239;
RN      11
RP      SEQUENCE FROM N.A.
RC      STRAIN=BRISTOL N2;
RX      MEDLINE=94150718; PubMed=7906398;
RA      Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA      Bradford J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
RA      Cretton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA      Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA      Jones M., Keshav J., Kirsten J., Laister N., Latreille P.,
RA      Lightnig J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA      Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA      Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA      Thierry-Mieg J., Thomas K., Vaubin M., Vaughan K., Waterston R.,
RA      Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,
RA      "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT      elegans."
RT      Nature 368:32-38(1994).
RN      12
RP      SEQUENCE FROM N.A.
RC      STRAIN=BRISTOL N2;
RX      Bradshaw H., Wohlmann P.;
RA      Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN      13
RP      SEQUENCE FROM N.A.
RC      STRAIN=BRISTOL N2;
RX      Waterston R.;
RA      Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN      14
RP      EMBL; U80455; AAB37887.1;
RX      SEQUENCE 411 AA; 44675 MW; 599DAC9DCFA81382 CRC64;

Query Match      36.8%; Score 92; DB 5; Length 411;
Best Local Similarity 39.5%; Pred. No. 0.00063;
Matches 17; Conservative 9; Mismatches 13; Indels 4; Gaps 2;
QY      2 QDDPOTDCQ--CORRCROESGPRQYQYCORRCKELCEEEERY 43
      1 111 11 111: 11 11111: 1: :111
Db      138 QQPAGQCCQDCQCAACPGQQ--QFQQCCQCCQCTTCGSSDDY 177

RESULT      7
P019594
ID      P019594      PRELIMINARY:      PRT:      572 AA.
AC      P019594:
DT      01-JUN-1998 (TREMBLrel. 06, Created)
DT      01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT      01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE      HYPOPHETICAL 62.4 KDA PROTEIN F19G12.7 IN CHROMOSOME X PRECURSOR.
GN      F19G12.7.
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC      Rhabditidae; Peloiderinae; Caenorhabditis.
OX      NCBI_TaxID=6239;
RN      11
RP      SEQUENCE FROM N.A.
RC      STRAIN=BRISTOL N2;
RX      Nhan M., Le T.T.;

```

```

RC Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
CL -1- SIMILARITY: BELONGS TO FAMILY UPF.
DR EMBL; U51997; AAC48159.1; -.
DR WORMPEP; F19612.7; CE07090.
KW Hypothetical protein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 572
FT SEQUENCE 572 AA; 62384 MW; BDA552FDB850AAB9 CRC64;
SO
Query Match
Best Local Similarity 31.2%; Score 78; DB 5; Length 572;
Matches 19; Conservative 7; Mismatches 7; Indels 20; Gaps 4;
QY 5 POTDQQQCCRRCKRQ-----QESGPRQDYCCRRCKE---ICEEE 40
DB 380 PQ-CQQCCQAPVCEQAPVCEQCCASLCHQSPAPCCQ-CQNTCCQAPVCEQO 429
RESULT 8
017400 PRELIMINARY; PRT; 425 AA.
AC 017400;
DT 01-JAN-1999 (TREMBLrel. 09, Created)
DT 01-JAN-1999 (TREMBLrel. 09, Last sequence update)
DT 01-JAN-1999 (TREMBLrel. 09, Last annotation update)
DE HYPOTHETICAL 45.9 KDA PROTEIN AC3.3 IN CHROMOSOME V PRECURSOR.
GN AC3.3
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA McMurtry A.;
RL Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
CC -1- SIMILARITY: BELONGS TO FAMILY UPF.
DR EMBL; Z71177; CAA94867.1; -.
DR WORMPEP; AC3.3; CE05133.
KW Hypothetical protein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 425
FT SEQUENCE 425 AA; 45922 MW; 951352A2AF7E96C CRC64;
SQ
Query Match
Best Local Similarity 31.0%; Score 77.5; DB 5; Length 425;
Matches 18; Conservative 9; Mismatches 7; Indels 11; Gaps 4;
QY 6 QYDCQ---QCRRCR--QESGPRQDYCCRRCKE---ICEEE 40
DB 236 QYTCQQAQVPCQCCQAPCCQPSAPCCQ-CQNTCCQQAQVPCQO 279
RESULT 9
017401 PRELIMINARY; PRT; 600 AA.
AC 017401;
DT 01-JAN-1999 (TREMBLrel. 09, Created)
DT 01-JAN-1999 (TREMBLrel. 09, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE HYPOTHETICAL 67.7 KDA PROTEIN AC3.4 IN CHROMOSOME V.
GN AC3.4
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA McMurtry A.;
RL Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.

```

CC	-1-	SUBCELLULAR LOCATION:	INTEGRAL MEMBRANE PROTEIN (POTENTIAL),
CC	-1-	SIMILARITY: BELONGS TO FAMILY UPF.	
DR	EMBL: Z71177;	CAA94868.1; -.	
DR	WORMPEP; AC3.4;	CE05134.	
DR	INTERPRO: IPR001594;	-.	
DR	PFAM: PF01529;	zf-DHHC; 1.	
DR	PRODOM: PD003041;	?; 1.	
KW	Hypothetical protein;	Transmembrane.	
FT	TRANSMEM	309	329
FT	TRANSMEM	447	467
FT	TRANSMEM	490	510
FT	DOMAIN	6	287
SO	SEQUENCE	600 AA;	67740 MW; 504A5C61BA72091B CRC64;

Query Match	31.0%;	Score 77.5;	DB 5;	Length 600;
Best Local Similarity	40.0%;	Pred. No. 0.051;		
Matches 18; Conservative	9;	Mismatches 7;	Indels 11;	Gaps 4;

```
QY      6 QTDCC---QCQRRCR---QESGPRQQYCQRCKE---ICEEE 40
          || ||      ||::|      || |      || |      ||::
Db     103 QTTCCQAVPQCQQQCAPQCQQPSAPQCQQ-CQNTCCQAAVPCQQ 146
```

RESULT	10	
09ZMW13		
ID	09ZMT3	PRELIMINARY:
AC	09ZMT3;	PRT; 810 AA.
DT	01-MAY-1999 (TREMBLrel. 10)	Created)
DT	01-MAY-1999 (TREMBLrel. 10)	Last sequence update)
DT	01-OCT-2000 (TREMBLrel. 15)	Last annotation update)
DE	PV100.	
OS	Cucurbita maxima (Pumpkin)	(Winter squash).
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;	
OC	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;	
OC	Cucurbitales; Cucurbitaceae; Cucurbita.	
OX	NCBI_TaxID=3661;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=KUROKAWA AMAKURI MANKIN; TISSUE=COYYLEDON;	
RC	MEDLINE=99107919; PubMed=9891029;	
RA	Yamada K., Shimada T., Kondo M., Nishimura M., Hara-Nishimura I.;	
RT	"Multiple functional proteins are produced by cleaving Asn-Gln bonds of a single precursor by vacuolar processing enzyme.";	
RL	J. Biol. Chem. 274:2563-2570(1999).	
DR	EMBL; AB019195; BAA34056.1; -.	
DR	HSSP; P02853; 2PRL.	
DR	INTERPRO; IPR001113; -.	
DR	PFAM; PF00546; Seedstore_7s; 1.	
DR	PRODOM; PD081059; -; 1	
SO	SEQUENCE	810 AA; 97314 MW; AB29A3F7542266AB CRC64;

Query Match	30.8%;	Score 77;	DB 10;	Length 810;
Best Local Similarity	34.28;	Pred. No. 0.076;		
Matches 13;	Conservative 10;	Mismatches 15;	Indels 0;	Gaps 0

Dy 5 PQTDCQQCQRRCRQDSGPRQOYCQRCKEICEEEEE 42  
| : : | : || - | : || : || : | : |  
Db 74 PRAEYEVCRLCQVAERGVGEQQRKCEQCCELRRLREQ 111

RESULT	11	
Q19919		
ID	Q19919	PRELIMINARY;
		PRT;
		242 AA

DT 01-JUN-1998 (TREMBLrel.06, Created)  
DT 01-JUN-1998 (TREMBLrel.06, Last sequence update)  
DT 01-JUN-1998 (TREMBLrel.06, Last annotation update)  
DE HYPOTHETICAL, 26.2 KDA PROTEIN F3IA3.1 IN CHROMOSOME X

OS *Caenorhabditis elegans*.  
OC Eukaryota; Metazoa; Chromadorea; Rhabditida; Rhabditoidea.

OC Rhabditiidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
OX

RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RA Murray J, Le T.T.;  
RL Submitted (May-1996) to the EMBL/GenBank/DBJ databases  
CC 1. SIMILARITY: BELONGS TO FAMILY UFP.

KM	Hypothetical protein; Transmembrane.
FT	TRANSMEM 3 23 POTENTIAL.
SQ	SEQUENCE 242 AA; 26213 MW; C7D6996BD0829317 CRC64;

Query Match	29.48;	Score 73.5;	DB 5;	length 242;
Best Local Similarity	33.38;	Pred. NO. 0.076;		
Matches 17;	Conservative 10;	Mismatches 13;	Indels 11;	Gaps 4;

```

Qy      2 QDDPQTDCQCQCRKR-----QQESGR-QQY---CQRCKEICEEEE 42
          |   |||  | :   || : ||| : || : || : :
Db      77 QQPAAPQCQCQNTCQGSAPVCQQQCAPIQQQQFAPACQ-QCQNSCQQTQQ 126

```

RESULT	12	
ID	Q19054	
AC	Q19054	PRELIMINARY; PRT: 111 AA.
DT	01-NOV-1996 (TREMBLrel. 01, Created)	
DT	01-JAN-1998 (TREMBLrel. 05, Last sequence update)	
DT	01-JAN-1999 (TREMBLrel. 09, Last annotation update)	
DE	E04D5.4 PROTEIN.	
GN	E04D5.4.	
OS	Caenorhabditis elegans.	
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;	
OC	Rhabditidae; Peloiderinae; Caenorhabditis.	
OX	NCBI_TaxID=6239;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	McMurray A.;	
RI	Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.	

RP SEQUENCE FROM N.A.  
RX MEDLINE=94150718; PubMed=7906398;  
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
RA Jones M., Kershaw J., Kirsten T., Laister N., Laetle P.,  
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
RA Parsons J., Percy C., Ritken L., Roopra A., Saunders D., Shownkeen R.,  
RA Shalton N., Smith A., Sonhammer E., Staden R., Sulston J.,  
RA Thiertry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterson R.,  
RA Watson A., Welstock L., Wilkinson-Spratt J., Wohlman P.,  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans".  
RL Nature 368:32-38(1994).  
DR EMBL: Z66496; CAA91281.1; -  
SQ SEQUENCE 111 AA; 11681 MW; AAF595F0DA154FF33 CRC64;

Query Match 29.2%; Score 73; DB 5; Length 111;  
Best Local Similarity 41.9%; Pred. No. 0.047;  
Matches 13; Conservative 3; Mismatches 15; Indels 0; Caps 0;

```

QY      3 DDPQTDCQCQCRKCRQQESGPRQOQYQCRRC 33
          ||| ||| | : ||| : |
DB      61 DDPNTDCTQYTFELCSNAKYTPLLQGFCAKTC 91

```

RESULT 13  
018118  
ID 018118 PRELIMINARY; PRT; 330 AA

```

AC 018118;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
DT T23FL.6 PROTEIN.
GN T23FL.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilkison J.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RX MEDLINE=94150718; PubMed=7906398;
SEQUENCE FROM N.A.
RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Smalton N., Smith A., Sonhammer E., Staden K., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Wellstock L., Wilkison-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
DR EMBL: Z81129; CAB03405.1; -.
SQ SEQUENCE 330 AA; 36605 MW; F043B1A90D3A8FE9 CRC64;

Query Match 28.8%; Score 72; DB 5; Length 330;
Best Local Similarity 29.5%; Pred. No. 0.15;
Matches 13; Conservative 10; Mismatches 9; Indels 12; Gaps 2;

QY 7 TDCQCCQRCRCRQ-----ESGPRQO---YCQRCKEICE 38
DB 230 TNCQCCQNSCQNSMTQITTYVQASPTSCVCVQCQCCQCCQCC 273

RESULT 14
Q9N4R1 PRELIMINARY; PRT; 339 AA.
AC Q9N4R1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE HYPOTHETICAL PROTEIN Y5H2A.C.
GN Y5H2A.C.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN-BRISTOL N2;
RX STRAIN-BRISTOL N2;
WA Waterston R.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC006809; AAF59626.1; -.
SQ SEQUENCE 339 AA; 36218 MW; 40233423237CA32D CRC64;

```

```

Query Match 28.8%; Score 72; DB 5; Length 339;
Best Local Similarity 32.7%; Pred. No. 0.15;
Matches 17; Conservative 9; Mismatches 8; Indels 18; Gaps 4;

QY 5 PQRDCQCCQRCRCR-----QESGPR-----QOYQCRCKE---ICEEE 40
DB 172 PQ--CCQCCQNTCCQFAPVCCQCCAPCCTTSSAPCCQCCQFAPVCCQ 221

RESULT 15
Q16500 PRELIMINARY; PRT; 388 AA.
AC Q16500;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE C03A7.4 PROTEIN.
GN C03A7.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Smalton N., Smith A., Sonhammer E., Staden K., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Wellstock L., Wilkison-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN-BRISTOL N2;
RX GRECO T., Bradshaw H., Elliott G.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA STRAIN-BRISTOL N2;
RX Waterston R.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF016451; AAB66001.1; -.
SQ SEQUENCE 388 AA; 42139 MW; 2E20655B0B9AE492 CRC64;

QY 6 QTDCCQ--QCQRCRC-----RQESGPRQOQYQCRCKEICEE 39
DB 194 QACQCPQCCQCCQCCSCVQCCQCCQCCQCCQCCQCCQCCQCCQCC 233

```

Search completed: March 1, 2001, 16:09:14  
 Job time: 1558 sec

